

# Dynamic Bayesian Nonlinear Calibration

Derick L. Rivers\* and Edward L. Boone

Department of Statistical Sciences and Operations Research, Virginia  
Commonwealth University

## Abstract

*Statistical calibration where the curve is nonlinear is important in many areas, such as analytical chemistry and radiometry. Especially in radiometry, instrument characteristics change over time, thus calibration is a process that must be conducted as long as the instrument is in use. We propose a dynamic Bayesian method to perform calibration in the presence of a curvilinear relationship between the reference measurements and the response variable. The dynamic calibration approach adequately derives time dependent calibration distributions in the presence of drifting regression parameters. The method is applied to simulated spectroscopy data based on work by Lundberg and de Maré (1980).*

## 1. INTRODUCTION

In many areas such as analytical chemistry, bioassay, spectroscopy, and radiometry fitting a curve through data to perform statistical calibration is of great importance. The statistical calibration problem is typically carried out in two stages; first samples are collected consisting of observations and known reference measurements of a targeted subject, and second a fitted curve is established from the first stage and an observed value  $y_0$  is used to predict an unknown targeted reference measurement  $x_0$ . The linear approach to this problem has been given much consideration from both the frequentist perspective (Eisenhart 1939; Krutchkoff 1967; Berkson 1969; Williams 1969; Halperin 1970; Martinelle 1970; Lwin and Maritz 1982) and the Bayesian perspective (Hoadley 1970; Hunter and Lamboy 1981; Eno 1999). The multivariate case to the linear calibration problem is considered by Brown (1982) from both perspectives. Bayesian dynamic approaches to the linear statistical calibration problem have been explored that consider calibration estimates as a function of time (Smith and Corbett 1987; Rivers and Boone 2014).

Unfortunately in many cases these curves are curvilinear, and straight-line linear methods are inappropriate. Several authors have considered Bayesian nonlinear approaches to the calibration problem. Racine-Poon (1988) used a Bayesian approach to a nonlinear calibration problem arising from agrochemical soil bioassays (Osborne 1991). Racine-Poon (1988) show that the posterior distribution of an unknown concentration  $\eta$  can be calculated by several methods: maximum likelihood; a numerical integration method based on the Gauss quadrature approach of Naylor and Smith (1982); or an approximation based on the Laplace method for integrals (Tierney and Kadane 1986). A noninformative reference prior (Bernardo 1979) approach for the polynomial calibration model is presented by Eno and Ye (2000). Through a second-degree bioassay example presented by Aitchison and Dunsmore (1975), Eno and Ye (2000) derive a reference prior and make posterior inferences about the calibration distribution. In cases when it is not feasible to transform the data to create a straight line, Eno and Ye (2000) show that the inclusion of a quadratic term

\*Corresponding author. Email: riversdl@vcu.edu

appropriately adds flexibility to the model. A Bayesian random effects model is proposed by Fong et al. (2012) for the nonlinear calibration problem. Fong et al. (2012) proposed a calibration method that is robust to dependent outliers. They demonstrated the proposed method on data from the HIV Vaccine Trials Network Laboratory and used a normal-mixture model with dependent error terms to model the experimental noise.

Graphite furnace atomic absorption spectroscopy (GFAAS) is an analytical technique for determining trace metal concentrations in different samples. In GFAAS, calibration curves tend to be nonlinear and polynomial regression is used to evaluate unknown sample concentrations. Given a known sample of concentrations ( $x_i$ ) and corresponding absorbance values ( $y_i$ ) the following relationship can be assumed:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i, \quad i = 1, 2, \dots, m, \quad (1)$$

where  $\beta_0, \beta_1, \beta_2$  are the unknown parameters and  $\epsilon_i$  is an independent zero-mean error term with variance  $\sigma^2$ . Also, an observed absorbance  $y_0$  corresponding to an unknown concentration  $\xi$  can be modeled as

$$y_0 = \beta_0 + \beta_1 \xi + \beta_2 \xi^2 + \epsilon, \quad (2)$$

where the error term again is assumed to be independent with mean of zero and variance  $\sigma^2$ . The estimate of the unknown value  $\xi$  is calculated as follow:

$$\xi^* = \frac{-\hat{\beta}_1 + \sqrt{\hat{\beta}_1^2 - 4\hat{\beta}_2(\hat{\beta}_0 - y_0)}}{2\hat{\beta}_2}, \quad (3)$$

where  $\hat{\beta}_0, \hat{\beta}_1$  and  $\hat{\beta}_2$  are the least squares estimates of  $\beta_0, \beta_1$  and  $\beta_2$ . The estimate of  $\xi$  is denoted as  $\xi^*$  and  $y_0$  is the observed value associated with the unknown concentration  $\xi$ . Using Equations (1) - (3) Lundberg and de Maré (1980) propose a simple interval estimation approach to the spectroscopy calibration problem when there is small measurement error.

For sufficiently small values of the measurement variance  $\sigma^2$ , Lundberg and de Maré (1980) state that there is a unique and consistent estimate  $\xi^*$  of  $\xi$  and an asymptotic confidence interval of  $\xi$  is obtained as

$$\xi^* \pm t_{\alpha/2, (m+n)} d(\xi^*)$$

where  $d(\xi^*)$  is chosen as

$$d(\xi^*) = \frac{s \left[ \frac{1}{n} + \Xi'(\mathbf{X}'\mathbf{X})^{-1}\Xi \right]^{1/2}}{\left| \sum_{k=1}^2 k \hat{\beta}_k \xi^{*k} \right|},$$

where  $s^2$  is the residual variance,  $\mathbf{X}$  is the design matrix and

$$\Xi = \begin{bmatrix} 1 \\ \xi^* \\ \xi^{*2} \end{bmatrix}.$$

François et al. (2004) examined optimal designs for linear and nonlinear calibration models. They show that for the quadratic model given by Equation (1), the design points should be to the left part of the calibration domain where the calibration curve slope is smaller. François et al. (2004) state that in this area of the domain, the calibration prediction variance is higher and the design points aim then at decreasing the lack of predictive ability of the model in that area. Conversely, the delta method is used to derive an asymptotic confidence interval for  $\xi$  by François

et al. (2004) as well as Kirkup and Mulholland (2004). The variance in  $\xi^*$ , written as  $\sigma_{\xi^*}^2$ , is given by

$$\sigma_{\xi^*}^2 = \left( \frac{\partial \xi^*}{\partial y_0} \sigma_{y_0} \right)^2 + \mathbf{d}_{\xi^*}' \mathbf{V} \mathbf{d}_{\xi^*}, \quad (4)$$

where  $\mathbf{V}$  is the variance-covariance matrix of  $\boldsymbol{\beta}$  and

$$\mathbf{d}_{\xi^*} = \begin{bmatrix} \frac{\partial \xi^*}{\partial \beta_0} \\ \frac{\partial \xi^*}{\partial \beta_1} \\ \frac{\partial \xi^*}{\partial \beta_2} \end{bmatrix},$$

thus deriving an asymptotic confidence interval for  $\xi$

$$\xi^* \pm z_{1-\alpha/2} \sqrt{\sigma_{\xi^*}^2}. \quad (5)$$

For an overview of linear and curvilinear calibration methods that commonly use the quadratic calibration model see Merkle (1983); Kirkup and Mulholland (2004); Lavagnini and Magno (2006); Lim and Yun (2010).

Weinreb et al. (1990) address nonlinearity in calibration of the advanced very high resolution radiometer (AVHRR). It is assumed that AVHRR can be calibrated by only two points; an internal calibration target (ICT) and space, but Weinreb et al. (1990) state that by not accounting for nonlinearity, errors as large as 2°C in inferred scene temperatures. The most direct way to handle the nonlinearity would be to use a quadratic calibration equation (Weinreb et al. 1990).

Calibrations are never concluded once and for all. Instrument characteristics are altered by time and use, especially in radiometry, and calibration must be viewed as an iterative process as long as the instrument is in use (Cervenka and Massa 1994). Our study is motivated by extending the dynamic linear calibration model of Rivers and Boone (2014) to incorporate a quadratic term in the presence of nonlinearity. In Section 2, we introduce a Bayesian dynamic nonlinear calibration model akin to that of Lundberg and de Maré (1980); Weinreb et al. (1990); Eno (1999); Eno and Ye (2000); François et al. (2004); Kirkup and Mulholland (2004); Hibbert (2006); Lavagnini and Magno (2006); Lim and Yun (2010). In Section 3, we demonstrate through a simulation study how the dynamic nonlinear calibration model performs alongside the static estimator given by Equation (3) under various noise conditions. In Section 4 the proposed method is applied to a spectroscopy example and a microwave radiometry example. In the first example, the method is used to determine trace amounts of cadmium (Cd) in water samples and for the second example, it is used to estimate a reference temperature given an observed voltage output measure. In Section 5 we conclude with future work and other considerations.

## 2. DYNAMIC NONLINEAR CALIBRATION MODEL

When collecting laboratory or field measurements for the purpose of calibration, scientist and engineers face a problem when the subsequent stability of the instrument change in relation to time, temperature, pressure, or some other external factors. These changes may cause the instrument readings to drift since an initial calibration thus making it necessary to recalibrate the instrument (Ziener and Strauss, 1978). We address this problem by developing a dynamic calibration approach that detects changes in the calibration constants in the presence of nonlinearity.

Let  $\{(\mathbf{X}, \mathbf{Y}_t) | t = 1, 2, \dots, T\}$  be the reference measurement and responses in the calibration

experiment at time  $t$  and suppose the relationship can be described by

$$\mathbf{Y}_t = \mathbf{X}\boldsymbol{\beta}_t + \boldsymbol{\epsilon}_t, \quad t = 1, 2, \dots, T, \quad (6)$$

where  $\mathbf{Y}_t$  is a series of  $r$ -dimensional vector of responses,  $\mathbf{X}$  is the fixed  $(r \times d)$  reference design matrix,

$$\begin{bmatrix} 1 & X_1 & X_1^2 \\ 1 & X_2 & X_2^2 \\ \vdots & \vdots & \vdots \\ 1 & X_r & X_r^2 \end{bmatrix},$$

$\boldsymbol{\beta}_t$ , (i.e.  $\beta_{0t}$ ,  $\beta_{1t}$ , and  $\beta_{2t}$ ), is a series of  $d$ -dimensional vectors of unknown dynamic regression parameters, and  $\boldsymbol{\epsilon}_t$  is a  $r$ -dimensional vector of independently normally distributed error terms with mean  $\mathbf{0}$  and variance-covariance matrix  $\mathbf{E} = \sigma_E^2 \mathbf{I}$ . Equation (6) is known as the observation equation.

The evolving relationship between  $\mathbf{X}$  and  $\mathbf{Y}_t$  is expressed by the dynamic parameter vector  $\boldsymbol{\beta}_t$ . The evolution in time of the regression parameters is modelled as

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\omega}_t, \quad t = 1, 2, \dots, T, \quad (7)$$

where  $\boldsymbol{\omega}_t$  is a  $d$ -dimensional vector of independently normally distributed error terms with mean  $\mathbf{0}$  and variance-covariance matrix  $\mathbf{W} = \sigma_W^2 \boldsymbol{\Omega}$ , and  $\boldsymbol{\Omega} = [\mathbf{X}'\mathbf{X}]^{-1}$ . Equation (7) is known as the system equation.

The observation vector  $\mathbf{Y}_t$  and the dynamic regression parameter vector  $\boldsymbol{\beta}_t$  are both random variables, thus the expected values  $\hat{\mathbf{Y}}_t$  and  $\hat{\boldsymbol{\beta}}_t$  are the means of their respective distributions and must be estimated sequentially. The one-step forecast for  $\mathbf{Y}_t$  and posterior distributions  $\boldsymbol{\beta}_t$  for each time  $t$  are as follows in Algorithm 1 (Dynamic Linear Regression Models algorithm). See West et al. (1985); West and Harrison (1997) for a more detailed discussion of Dynamic Linear Regression Models (DLRMs).

Furthermore, let  $\{y_{0t}|t = 1, 2, \dots, T\}$  be the observation from the second stage of the calibration experiment corresponding to an unknown reference of interest  $x_{0t}$ , and

$$y_{0t} = \beta_{0t} + \beta_{1t}x_{0t} + \beta_{2t}x_{0t}^2 + \epsilon_{0t}, \quad t = 1, 2, \dots, T, \quad (8)$$

where  $\beta_{0t}$ ,  $\beta_{1t}$ , and  $\beta_{2t}$  are the time dependent regression coefficients and  $\epsilon_{0t}$  are assumed to be independently normally distributed with mean 0 and variance  $\sigma_E^2$ . Since the quadratic model in Equations (6) and (7) are not monotonic on  $\mathbb{R}$ , the domain will be restricted to where it is strictly increasing.

We assume that the first stage of calibration experiment is independent of the second stage, therefore  $x_{0t}$  is independent of  $(\boldsymbol{\beta}_t, \boldsymbol{\Gamma})$  where the joint prior distribution is

$$\pi(x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma}) = \pi(x_{0t})\pi(\boldsymbol{\beta}_t, \boldsymbol{\Gamma}) \quad (9)$$

and  $\boldsymbol{\Gamma}' = [\sigma_E^2 \quad \sigma_W^2]$ . The posterior of  $(x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma})$  is then given by

$$\begin{aligned} \pi(x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma}|y_{0t}, \mathbf{Y}_t) &\propto f(y_{0t}, \mathbf{Y}_t|x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma})\pi(x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma}) \\ &\propto f(y_{0t}|x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma})\pi(x_{0t})f(\mathbf{Y}_t|\boldsymbol{\beta}_t, \boldsymbol{\Gamma})\pi(\boldsymbol{\beta}_t, \boldsymbol{\Gamma}) \\ &\propto f(y_{0t}|x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma})\pi(\boldsymbol{\beta}_t, \boldsymbol{\Gamma}|\mathbf{Y}_t)\pi(x_{0t}), \end{aligned} \quad (10)$$

---

**Algorithm 1:** Updating (DLRM) Dynamic Linear Regression Model
 

---

```

Initialize  $t = 0$ 
{Initial information  $(\beta_0|D_0) \sim N_d[\mathbf{m}_0, \mathbf{C}_0]$ }
Input:  $\mathbf{m}_0, \mathbf{C}_0, \mathbf{E}, \mathbf{W}$ 
loop
   $t = t + 1$ 
  {Compute prior at t:  $(\beta_t|D_{t-1}) \sim N_d[\mathbf{a}_t, \mathbf{R}_t]$ }
     $\mathbf{a}_t = \mathbf{m}_{t-1}$ 
     $\mathbf{R}_t = \mathbf{C}_{t-1} + \mathbf{W}$ 
  Input:  $\mathbf{X}$ 
  {Compute forecast at t:  $(\mathbf{Y}_t|D_{t-1}) \sim N_r[\mathbf{f}_t, \mathbf{Q}_t]$ }
     $\mathbf{f}_t = \mathbf{X}\mathbf{a}_t$ 
     $\mathbf{Q}_t = \mathbf{X}\mathbf{R}_t\mathbf{X}' + \mathbf{E}$ 
  Input:  $\mathbf{Y}_t$ 
  {Compute forecast error  $\mathbf{e}_t$ }
     $\mathbf{e}_t = \mathbf{Y}_t - \mathbf{f}_t$ 
  {Compute adaptive gain matrix  $\mathbf{A}_t$ }
     $\mathbf{A}_t = \mathbf{Q}_t^{-1}\mathbf{X}\mathbf{R}_t$ 
  {Compute posterior at t:  $(\beta_t|D_t) \sim N_d[\mathbf{m}_t, \mathbf{C}_t]$ }
     $\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t\mathbf{e}_t$ 
     $\mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t'\mathbf{Q}_t\mathbf{A}_t$ 
end loop
    
```

---

where  $\mathbf{Y}_t$  and  $y_{0t}$  are respectively the observations from the first and second stages of calibration. Our knowledge about the evolving relationship established in the calibration experiment at each time point is given by the posterior density  $\pi(\beta_t, \Gamma|\mathbf{Y}_t)$  which is the middle term in Equation (10) and found by Algorithm 1 for given values of  $\Gamma$ . Following Algorithm 1 and using multivariate normal theory (West and Harrison 1997) we have

$$\pi(\beta_t|\mathbf{Y}_t, \Gamma) \sim N_d(\mathbf{m}_t, \mathbf{C}_t),$$

where  $\mathbf{m}_t$  is the posterior mean and  $\mathbf{C}_t$  is the variance-covariance matrix of  $\beta_t$  at time  $t$ .

The first term in Equation (10),  $f(y_{0t}|x_{0t}, \beta_t, \Gamma)$ , is the likelihood function for the second stage

---

**Algorithm 2:** (SIR) Sampling Importance Resampling
 

---

1. Draw **sample** candidates  $(\Theta^{(1)}), \dots, (\Theta^{(m)})$  *i.i.d.* from  $g(\Theta)$
  2. Calculate the standardized **importance** weights,  $w(\Theta^{(i)}) = \frac{f(\Theta^{(i)})/g(\Theta^{(i)})}{\sum f(\Theta^{(i)})/g(\Theta^{(i)})}$  for  $i = 1, \dots, m$
  3. **Resample**  $\Gamma^{(1)}, \dots, \Gamma^{(n)}$  from  $\Theta^{(1)}, \dots, \Theta^{(m)}$  with replacement of probability  $w(\Theta^{(1)}), \dots, w(\Theta^{(m)})$  respectively.
- 

of the calibration experiment which provides information from the data and  $\pi(x_{0t})$  is the prior density for the unknown calibration reference  $x_{0t}$ . We wish to obtain the conditional posterior density  $\pi(x_{0t}|y_{0t}, \mathbf{Y}_t)$  at each time  $t$ . In order to achieve this we will have to integrate over  $(\beta, \Gamma)$ .

We reduce the parameter space by centering and scaling the data such that

$$\sum_{i=1}^r x_i = 0 \text{ and } \frac{1}{n} \sum_{i=1}^r x_i^2 = 1$$

and the intercept term  $\beta_{0t}$  is eliminated by moving the origin of the calibration (Hibbert 2006) so the model in Equation (8) is written as

$$y_{0t} - \bar{y}_t = \beta_{1t}(x_{0t} - \bar{x}) - \beta_{2t}(x_{0t}^2 - \bar{x}^2) + \epsilon_{0t}, \quad t = 1, 2, \dots, T, \quad (11)$$

where  $\bar{x} = \frac{1}{r} \sum_{i=1}^r x_i$  and  $\bar{x}^2 = \frac{1}{r} \sum_{i=1}^r x_i^2$ . In the second stage of the calibration experiment, an observation  $y_{0t}$  allows calculation of an unknown reference at time  $t$  by:

$$\hat{x}_{0t} = \frac{-\hat{\beta}_{1t} \pm \sqrt{\hat{\beta}_{1t}^2 - 4\hat{\beta}_{2t}(\bar{y}_t - y_{0t} - \hat{\beta}_{1t}\bar{x} - \hat{\beta}_{2t}\bar{x}^2)}}{2\hat{\beta}_{2t}}. \quad (12)$$

The quadratic model in Equation (12) has two possible roots on  $\mathbb{R}$  but the solution of interest depends on the sign of  $\hat{\beta}_{1t}$ . The solution is the increasing part of Equation (12) when  $\hat{\beta}_{1t} < 0$  and the decreasing part of Equation (12) is the solution when  $\hat{\beta}_{1t} > 0$ .

Given the reduced parameter space and Equation (11), the likelihood function at time  $t$  is expressed as

$$f(y_{0t}|x_{0t}, \beta_t, \Gamma) \propto \exp \left\{ -\frac{1}{2} \left[ \sigma_{Y_t}^{-2} (y_{0t} - \bar{y}_t - \beta_{1t}(x_{0t} - \bar{x}) - \beta_{2t}(x_{0t}^2 - \bar{x}^2))^2 \right] \right\},$$

the prior density at time  $t$  is

$$\pi(x_{0t}) \propto \exp \left\{ -\frac{x_{0t}^2}{2} \right\},$$

and integrating Equation (10) with respect to  $\beta_t$  for a given  $\Gamma$ , produces

$$\begin{aligned} \pi(x_{0t}|y_{0t}, \mathbf{Y}_t, \Gamma) &\propto f(y_{0t}|x_{0t}, \Gamma, \mathbf{Y}_t) \pi(x_{0t}) \\ &\propto \exp \left\{ -\frac{1}{2} \left[ \sigma_{Y_t}^{-2} (y_{0t} - \bar{y}_t - \beta_{1t}(x_{0t} - \bar{x}) - \beta_{2t}(x_{0t}^2 - \bar{x}^2))^2 \right] - \frac{1}{2} x_{0t}^2 \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \left[ \sigma_{Y_t}^{-2} (\hat{x}_{0t} - x_{0t})^2 \right] - \frac{1}{2} x_{0t}^2 \right\}. \end{aligned}$$

By completing the square, the posterior density for the unknown reference measurement at time  $t$  is

$$\pi(x_{0t}|y_{0t}, \mathbf{Y}_t, \Gamma) \sim N(\mu_{x_{0t}}, \sigma_{x_{0t}}^2), \quad (13)$$

with

$$\begin{aligned} \mu_{x_{0t}} &= \frac{x_{0t}}{1 + \sigma_{Y_t}^2}, \\ \sigma_{x_{0t}}^2 &= \frac{1}{1 + \sigma_{Y_t}^2}, \end{aligned}$$

and

$$\sigma_{Y_t}^2 = \text{tr}(\mathbf{Q}_t)$$

where  $\text{tr}(\cdot)$  denotes trace of the one-step forecast variance-covariance matrix in Algorithm 1.

To obtain posterior samples of Equation (13) for a given  $\Gamma$  we use the Sampling Importance Resampling (SIR) algorithm (Rubin 1987; Smith and Gelfand 1992; Givens and Hoeting 2005; Albert 2007) to draw random samples from the target distribution  $\pi(x_{0t}|y_{0t}, \mathbf{Y}_t, \Gamma)$  by using a candidate distribution  $g(\Theta)$  for  $\Gamma$ , thus deriving the posterior densities of interest. Algorithm 2 is the Sampling Importance Resampling algorithm.

For most applications, it is believed *a priori* that the observational variance  $\sigma_E^2$  is greater in magnitude than the system variance  $\sigma_W^2$ , such that

$$\sigma_W^2 < \sigma_E^2.$$

To enforce this belief about the variance relationship we utilize the following prior distributions:

$$\sigma_E^2 \sim \text{Uniform}(0, \alpha_E) \quad (14)$$

$$\sigma_W^2 | \sigma_E^2 \sim \text{Uniform}(0, \sigma_E^2), \quad (15)$$

where prior distributions (14) and (15) ensure the system variance to be less than the observation variance. Since these are proper prior distributions the resulting posterior distribution will also be proper. We combine prior distributions (14), (15), Algorithm 1 (DLM), Equation (13), and Algorithm 2 (SIR) together and propose the Dynamic Calibration Method in Algorithm 3.

---

**Algorithm 3:** Dynamic Calibration

---

1. Draw  $M$  *i.i.d.* **sample** candidates for  $(\sigma_E^2, \sigma_W^2)$  from  $g(\Theta)$ .
2. Calibration data are fit using the DLRM framework for each of the  $M$  proposal samples  $(\sigma_E^{2(m)}, \sigma_W^{2(m)})$ , with the prior moments for  $(\beta_0 | D_0)$  as  $\mathbf{m}_0 = \mathbf{1}_d$  and  $\mathbf{C}_0 = \mathbf{100I}_{(d \times d)}$ , where  $\mathbf{1}_d$  is a  $d$ -dimensional vector of ones:
  - a. Data are scaled and centered such that  $\sum_{i=1}^r x_i = 0$ ,  $\frac{1}{n} \sum_{i=1}^r x_i^2 = 1$  and  $\beta_{0t} = 0$ ;
  - b. Estimate  $\beta_t^{(m)} | \sigma_E^{2(m)}, \sigma_W^{2(m)}$  for the  $m^{\text{th}}$  proposal sample  $\Theta^{(m)}$  for all time  $t$ ;
  - c. Draw a sample from  $\pi(x_{0t} | y_{0t}, \mathbf{Y}_t, \Gamma)$  given the  $m^{\text{th}}$  proposal sample  $\Theta^{(m)}$  at time  $t$ ;
  - d. Calculate log-likelihoods,  $\log[f(\Theta^{(m)})]$ , for each  $(\sigma_{E_t}^{2(m)}, \sigma_{W_t}^{2(m)})$  pair.
3. Sampling Importance Resampling (SIR) is used to simulate samples of  $x_{0t} | \beta_t, \sigma_E^2, \sigma_W^2$  by accepting a subset of  $N$  from the proposal density to be distributed according to the posterior density  $\pi(\Gamma | \mathbf{Y}_t)$  with candidate density  $g(\Theta)$ .
  - a. Calculate the standardized **importance** weights,  $w(\Theta^{(1)}), \dots, w(\Theta^{(M)})$ , where  $w(\Theta^{(m)}) = \log[f(\Theta^{(m)})] - \log[g(\Theta^{(m)})]$  for the  $m^{\text{th}}$  proposal sample;
  - b. **Resample**  $N$  calibrated time series from the  $M$  proposal values with replacement given probabilities  $p(\Theta^{(m)})$  where

$$p(\Theta^{(m)}) = \frac{e^{w(\Theta^{(m)})}}{\sum_{j=1}^M e^{w(\Theta^{(j)})}}.$$

4. Rescale calibrated time series to original scale and record summary statistics (i.e. medians and credible sets) across each time  $t$ .
-

### 3. SIMULATION STUDY

In this section, we conducted several simulation experiments to illustrate the adaptability of the proposed method. In the simulation, we consider nine cases based on variance combinations for the observation and system variances,  $(\sigma_E^2, \sigma_W^2)$ , for the Dynamic Linear Models. The true values for  $\sigma_E^2$  and  $\sigma_W^2$  used in the simulation study are  $(0.00001, 0.0001, 0.001)$  and  $(0.00005, 0.0001, 0.001)$ , respectively. For each variance pair  $(\sigma_E^2, \sigma_W^2)$ , the number of simulated realizations is  $N = 100$ . In each realization, the number of simulated time periods is  $T = 1000$ . The posterior densities samples drawn from  $\pi(x_{0t}|y_{0t}, \mathbf{Y}_t, \mathbf{\Gamma})$  by the Dynamic Calibration Approach (Algorithm 3) is assessed via the square root of the averaged mean squared errors (*RAMSE*), the average interval width (*AvIW*), and the average coverage probability (*AvCP*):

$$RAMSE = \left[ \frac{1}{N} \sum_{j=1}^N MSE_j \right]^{\frac{1}{2}},$$

where

$$MSE_j = \frac{1}{T} \sum_{t=1}^T (\hat{x}_{0t} - x_{0t})^2;$$

and

$$AvIW = \frac{1}{N} \sum_{j=1}^N IW_j.$$

where

$$IW_j = \frac{1}{T} \sum_{t=1}^T (x_{0t}^U - x_{0t}^L).$$

Note that if  $x_{0t}^L$  is the 0.025 posterior quantile for  $x_{0t}$ , and  $x_{0t}^U$  is the 0.975 posterior quantile for  $x_{0t}$ , where  $x_{0t}$  is the true value of the calibration target from the second stage of experimentation, then  $(x_{0t}^L, x_{0t}^U)$  is a 95% credible interval.

Using the credible interval above we defined the covered probability (*CP*) which is calculated as such

$$CP_j = \frac{1}{T} \sum_{t=1}^T \psi_t$$

where

$$\psi_t = P[x_{0t}^L < x_{0t} < x_{0t}^U] = \begin{cases} 0 & \text{if } x_{0t} \notin (x_{0t}^L, x_{0t}^U); \\ 1 & \text{if } x_{0t} \in (x_{0t}^L, x_{0t}^U). \end{cases}$$

The average coverage probability (*AvCP*) is calculated by averaging across the number of replications in the simulation study, where

$$AvCP = \frac{1}{N} \sum_{j=1}^N CP_j.$$

The data are generated from the following model:

$$\mathbf{Y}_t = \mathbf{X}\boldsymbol{\beta}_t + \boldsymbol{\epsilon}_t, \tag{16}$$



where  $\mathbf{X}$  is a known fixed model matrix of reference values augmented with a column of 1's. The dynamic Bayesian nonlinear calibration model will be assessed across three different reference measurement schemes. In the first scheme, the reference measurements will be taken at [20, 90, 100]. The references will be placed at [20, 60, 90, 100] for the second scheme and at [20, 40, 60, 90, 100] for the final and third scheme. See Figure 1.

The vector of regression parameters,  $\beta_t$ , are random draws from a multivariate normal distribution with mean vector  $[\beta_0 \ \beta_1 \ \beta_2]'$  and variance-covariance matrix,  $\mathbf{W} = \sigma_W^2 [\mathbf{X}'\mathbf{X}]^{-1}$  for  $t = 1, \dots, T$ , where  $\beta_0 = -0.0007$ ,  $\beta_1 = 0.01858$ , and  $\beta_2 = -0.000117$ . For each  $t$ , the random multivariate error vector is

$$\epsilon_t \sim N_r[\mathbf{0}, \sigma_E^2 \mathbf{I}]. \quad (17)$$

Tables 1 - 3 summarizes the results of the dynamic nonlinear calibration estimator under different variance pairs,  $(\sigma_E^2, \sigma_W^2)$ , and results for the static model given by Equation (3). An asymptotic variance and 95% confidence interval for the static quadratic calibration is calculated from Equations (1) and (1). The dynamic calibration approach is denoted as DC and the static calibration approach is denoted as SC in the tables. From Tables 1 - 3, we make the following observations:

- In Table 1, the 3 reference case, the static method require that the degrees of freedom  $n - 3$  be  $\geq 1$  in order to estimate the variance  $\sigma^2$ . The dynamic calibration method is not dependent upon the degrees of freedom for estimation of the error variance, thus, credible intervals and interval widths can be computed.
- The RAMSE for the dynamic method is 11% to 41% smaller than the the RAMSE for the static method.
- The RAMSEs of the dynamic estimator and the static estimator increase as the observation variance  $\sigma_E^2$  increases.
- For the 4 and 5 reference models, when  $\sigma_E^2 = 0.001$  the average coverage propability (ACP) is just barely above 40%.
- Given  $\sigma_E^2$ , the RAMSEs increase slowly as  $\sigma_W^2$  increases.
- For the 4 and 5 reference models, the average interval width (AIW) increases for the dynamic method as the observation variance  $\sigma_E^2$  increase. The interval widths for the static calibration method is consistent regardless of the observational or system noise.

**Table 1:** Summary of the RAMSE, AIW and AvCP with 3 references for the dynamic nonlinear calibration estimator (DC) and the static quadratic estimator (SC) under  $(\sigma_E^2, \sigma_W^2)$ .

3 Reference Model						
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.00005$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	0.249	0.382	1.471	N/A	0.992	N/A
0.0001	0.643	1.099	4.642	N/A	0.997	N/A
0.001	1.988	3.391	14.592	N/A	0.997	N/A
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.0001$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	0.291	0.413	1.473	N/A	0.984	N/A
0.0001	0.661	1.109	4.645	N/A	0.996	N/A
0.001	1.999	3.399	14.592	N/A	0.997	N/A
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.001$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	0.716	0.806	1.500	N/A	0.688	N/A
0.0001	0.936	1.306	4.652	N/A	0.983	N/A
0.001	2.126	3.475	14.600	N/A	0.996	N/A

**Table 2:** Summary of the RAMSE, AIW and AvCP with 4 references for the dynamic nonlinear calibration estimator (DC) and the static quadratic estimator (SC) under  $(\sigma_E^2, \sigma_W^2)$ .

4 Reference Model						
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.00005$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	0.247	0.358	1.500	3.763	0.993	0.960
0.0001	0.739	1.072	4.741	3.920	0.994	0.931
0.001	2.332	3.376	14.875	3.920	0.995	0.451
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.0001$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	0.259	0.365	1.500	3.763	0.991	0.960
0.0001	0.747	1.081	4.742	3.920	0.994	0.930
0.001	2.336	3.379	14.875	3.920	0.994	0.451
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.001$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	0.427	0.498	1.502	3.920	0.917	0.998
0.0001	0.825	1.131	4.742	3.920	0.990	0.924
0.001	2.380	3.411	14.876	3.920	0.994	0.446

**Table 3:** Summary of the RAMSE, AIW and AvCP with 5 references for the dynamic nonlinear calibration estimator (DC) and the static quadratic estimator (SC) under  $(\sigma_E^2, \sigma_W^2)$ .

5 Reference Model						
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.00005$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	0.249	0.350	1.569	3.920	0.994	1.000
0.0001	0.768	1.083	4.955	3.920	0.995	0.927
0.001	2.429	3.374	15.515	3.920	0.995	0.452
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.0001$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	0.257	0.356	1.569	3.920	0.993	1.000
0.0001	0.770	1.082	4.955	3.920	0.995	0.929
0.001	2.432	3.377	15.515	3.878	0.995	0.447
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.001$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	0.370	0.445	1.570	3.920	0.958	0.999
0.0001	0.821	1.120	4.956	3.920	0.993	0.924
0.001	2.462	3.404	15.516	3.920	0.994	0.448

# 1. Simulation with Random Shock

**Table 4:** Summary of the RAMSE, AIW and AvCP with 3 references for the dynamic nonlinear calibration estimator (DC) and the static quadratic estimator (SC) under  $(\sigma_E^2, \sigma_W^2)$  with random shocks in system.

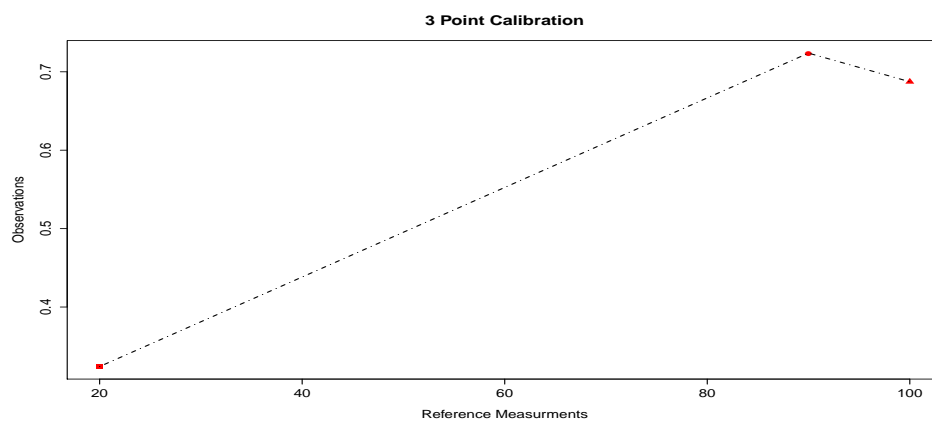
3 Reference Model						
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.00005$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	2.554	2.603	1.643	N/A	0.953	N/A
0.0001	2.587	2.788	4.770	N/A	0.958	N/A
0.001	3.115	4.234	14.652	N/A	0.966	N/A
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.0001$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	2.556	2.605	1.644	N/A	0.945	N/A
0.0001	2.589	2.790	4.770	N/A	0.958	N/A
0.001	3.120	4.239	14.653	N/A	0.966	N/A
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.001$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	2.619	2.674	1.672	N/A	0.674	N/A
0.0001	2.654	2.855	4.778	N/A	0.945	N/A
0.001	3.192	4.296	14.656	N/A	0.962	N/A

**Table 5:** Summary of the RAMSE, AIW and AvCP with 4 references for the dynamic nonlinear calibration estimator (DC) and the static quadratic estimator (SC) under  $(\sigma_E^2, \sigma_W^2)$  with random shocks in system.

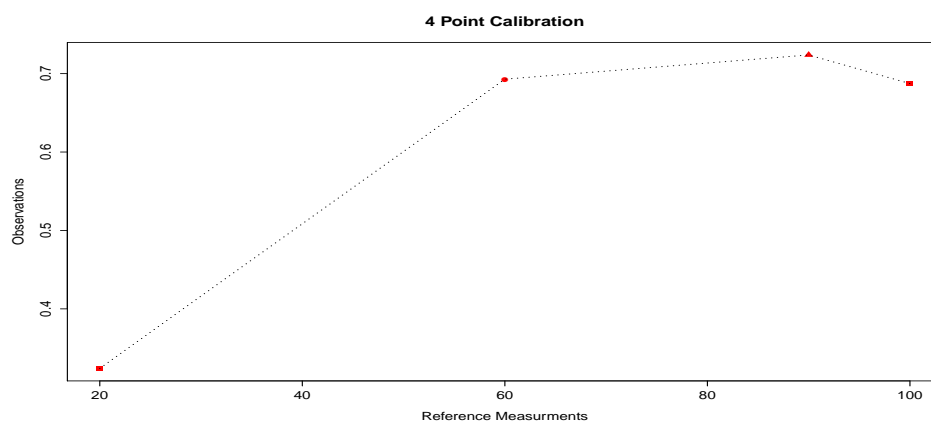
4 Reference Model						
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.00005$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	2.506	2.604	1.641	3.920	0.953	0.960
0.0001	2.562	2.789	4.844	3.920	0.956	0.889
0.001	3.301	4.231	14.921	3.920	0.967	0.429
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.0001$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	2.507	2.605	1.640	3.920	0.951	0.960
0.0001	2.563	2.790	4.844	3.920	0.956	0.889
0.001	3.303	4.234	14.921	3.920	0.967	0.429
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.001$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	2.524	2.621	1.642	3.920	0.880	0.959
0.0001	2.582	2.807	4.845	3.920	0.952	0.882
0.001	3.330	4.259	14.922	3.920	0.966	0.425

**Table 6:** Summary of the RAMSE, AIW and AvCP with 5 references for the dynamic nonlinear calibration estimator (DC) and the static quadratic estimator (SC) under  $(\sigma_E^2, \sigma_W^2)$  with random shocks in system.

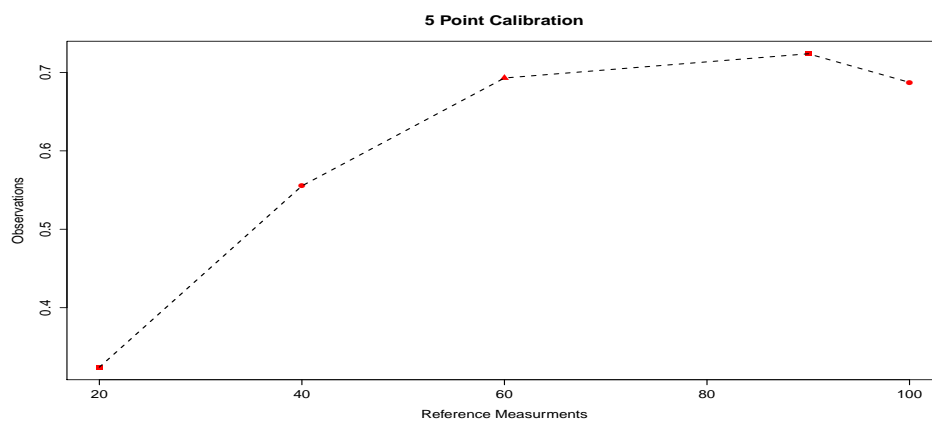
5 Reference Model						
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.00005$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	0.249	0.350	1.569	3.920	0.994	1.000
0.0001	0.768	1.083	4.955	3.920	0.995	0.927
0.001	2.429	3.374	15.515	3.920	0.995	0.452
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.0001$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	0.257	0.356	1.569	3.920	0.993	1.000
0.0001	0.770	1.082	4.955	3.920	0.995	0.929
0.001	2.432	3.377	15.515	3.878	0.995	0.447
$\sigma_E^2$	RAMSE		$\sigma_W^2 = 0.001$ AIW		ACP	
	DC	SC	DC	SC	DC	SC
0.00001	0.370	0.445	1.570	3.920	0.958	0.999
0.0001	0.821	1.120	4.956	3.920	0.993	0.924
0.001	2.462	3.404	15.516	3.920	0.994	0.448



**(a) 3-Point Calibration**



**(b) 4-Point Calibration**



**(c) 5-Point Calibration**

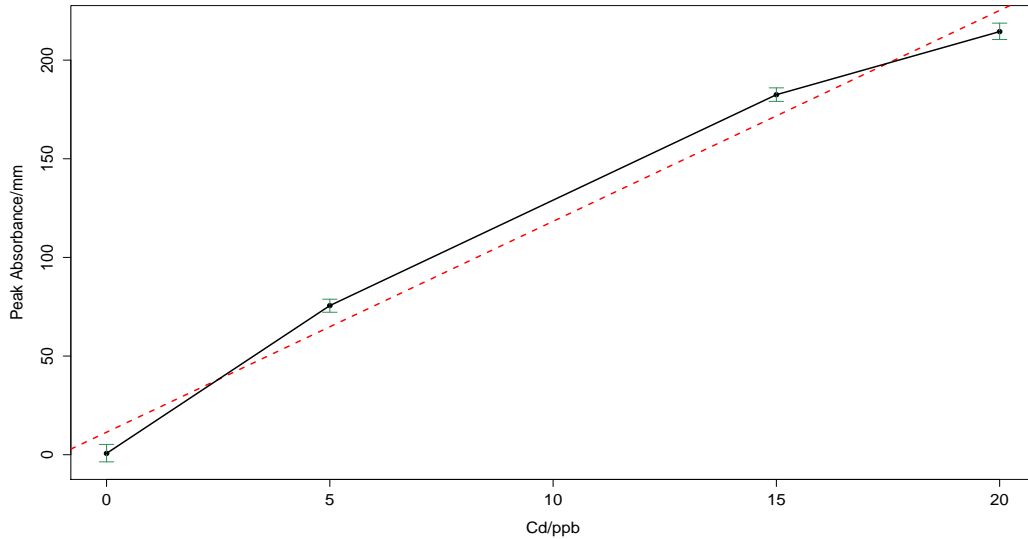
**Figure 1:** Three reference design schemes: (a) 3-point calibration; (b) 4-point calibration; (c) 5-point calibration



## 4. APPLICATIONS

### Example 1

As a demonstration of the dynamic nonlinear calibration method, we extend the example of nonlinear calibration presented by Lundberg and de Maré (1980). In Lundberg and de Maré's



**Figure 2:** Linear (dotted) and quadratic calibration curves for Cd using four standard references. Indicated are 95% credible intervals of the estimated absorbances.

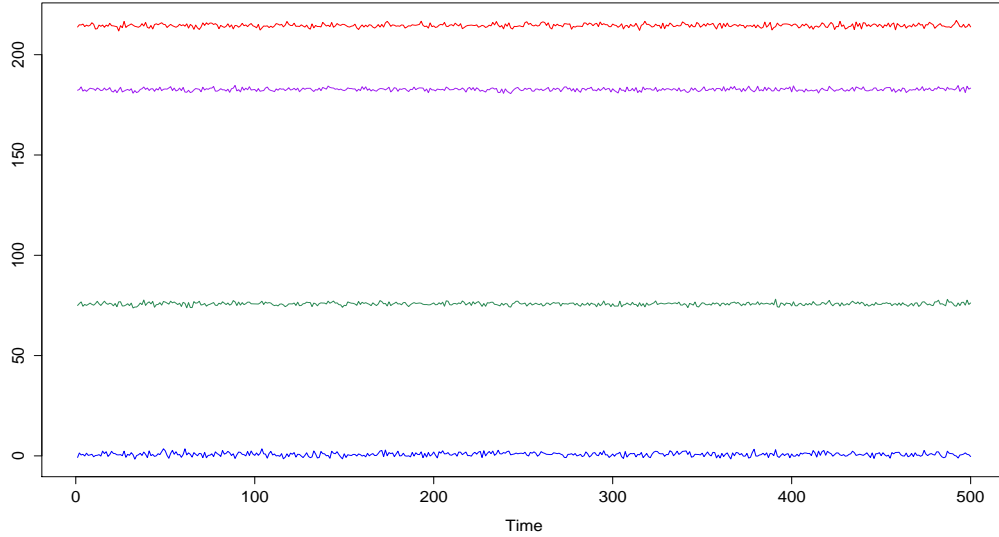
(1980) example graphite furnace atomic absorption spectroscopy (GFAAS) was used to determine trace amounts of cadmium (Cd) in water samples. In the first stage of experimentation two  $\mu\text{l}$  volumes of standard solutions containing 0, 5, 10, 15, and 20 ppb (parts per billion) of Cd were injected into a graphite furnace (Varian Techtron AA-6 spectrophotometer supplied with a Carbon Rod Atomizer model 63). The transient absorbance signals obtained when atomizing the standards

**Table 7:** Cd signals obtained when atomizing standards

Concentration/ppb	Peak Absorbance/mm					
0	0	1	1	0	1	
5	74	74	78	78	76	
15	183	184	178	183	184	
20	217	215	213	218	210	215

were recorded with a stripchart recorder. Each standard was run several times and a plot of peak absorbance (in mm) vs. concentration was made using the data in Table 7 (See Figure 2). At the second stage of experimentation the fifth reference (10 ppb) was used as an unknown sample to test the validity of the method with peak absorbance measurements of 135, 142, 132, 141, and 136. Lundberg and de Maré (1980) report an approximate 95% confidence interval for the unknown concentration with 10 being the true value as [9.7, 10.3].

We extend Lundberg and de Maré's (1980) example by generating 500 simulated peak



**Figure 3:** Peak Absorbance measurements: Concentration 0ppb (blue); Concentration 5ppb (green); Concentration 15ppb (violet); and Concentration 20ppb (red).

absorbance measurements for the four standard solutions of Cd (see Figure 3). The repeated measures for the peak absorbance were generated given a multivariate normal distribution  $\beta_t$  with mean vector

$$\mu = \begin{bmatrix} 0.72 \\ 16.448 \\ -0.288 \end{bmatrix}$$

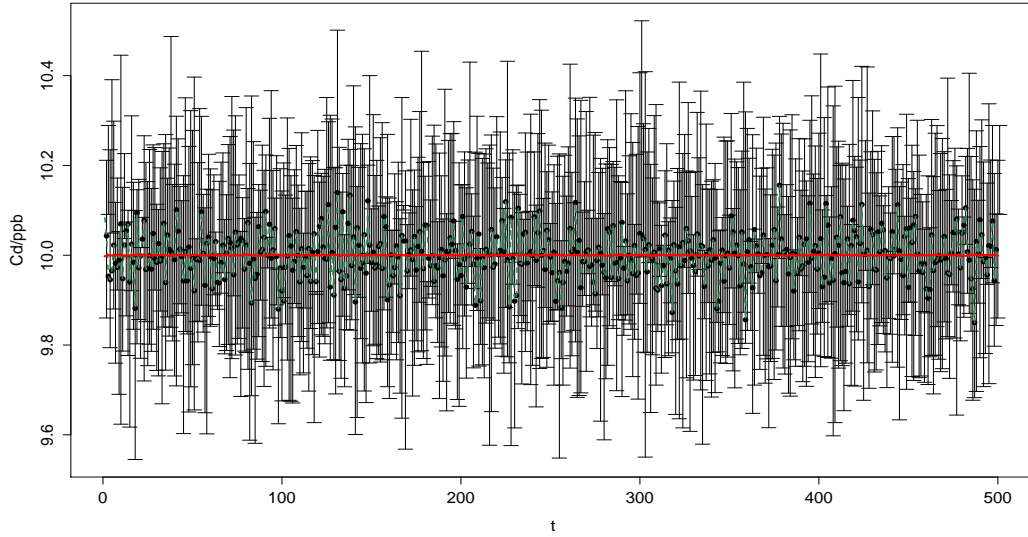
and variance-covariance matrix

$$\Sigma = \sigma^2 \begin{bmatrix} 0.17966 & -0.03435 & 0.00131 \\ -0.03435 & 0.01473 & -0.00069 \\ 0.00131 & -0.00069 & 0.00003 \end{bmatrix},$$

where  $\sigma^2 = 4.7$ , the residual variance from the ordinary least squares fit of the original data. The time series of the posterior estimates for Cd is given in Figure 4. The square root of the mean squared error is 0.0025 with an average 95% credible interval [9.8 10.2] which is shorter than the 95% confidence interval of [9.7 10.3] reported by Lundberg and de Maré (1980).

## Example 2

We turn to an example in microwave radiometry to also demonstrate the usefulness of the proposed dynamic calibration approach. Engineers and scientist commonly use microwave radiometers to measure electromagnetic radiation. This radiant power is emitted by some source or a particular surface such as ice or land surface. Radiometers are very sensitive instruments that are capable



**Figure 4:** Time series of Cd estimates (black dots connected by green lines) with corresponding 95% credible interval at time  $t$ . The "true" Cd of 10 ppb is given by the red line.

of measuring extremely low levels of radiation. The transmission source of the radiant power is the target of the radiometers antenna. When a scene, such as terrain, is observed by a microwave radiometer, the radiation received by the antenna is partly due to self-emission by the scene and partly due to the reflected radiation originating from the surroundings (Ulaby *et al.* 1981). This source may be cosmic background radiation, ocean surface, or a heated surface used for the purpose of calibration.

Calibration is required due to the fact that the current electronic hardware is unable to maintain a stable input/output relationship. For space observing instruments, stable calibration without any drifts is a key to detect proper trends of climate (Imaoka *et al.* 2010). Due to problems such as amplifier gain instability and exterior temperature variations of critical components that may cause this relationship to drift over time (Bremer 1979). During the calibration process, the radiometer receiver measures the voltage output power  $v(t)$ , and its corresponding input temperature of a known reference. Two or more known reference temperatures are needed for calibration of a radiometer. Ulaby *et al.* (1981); Racette and Lang (2005) state that the relationship between the output,  $v(t)$  and the input,  $T_A$  is approximately linear, and can be expressed as

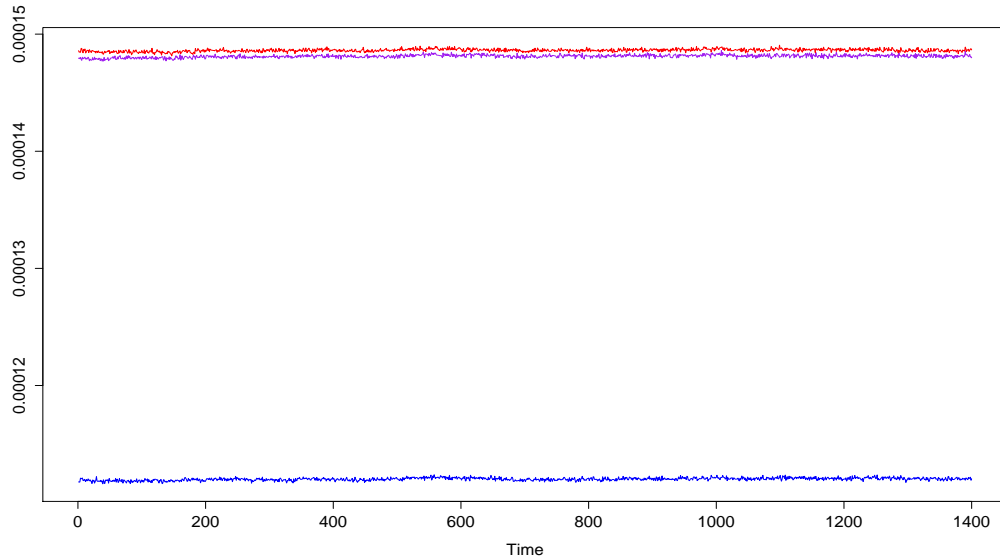
$$\hat{T}_A = \hat{\beta}_0 + \hat{\beta}_1 v(t)$$

where,  $\hat{\beta}_0$  and  $\hat{\beta}_1$  are the least square estimates for the regression parameters,  $\hat{T}_A$  is the estimated value of the brightness temperature and  $v(t)$  is the observed output voltage. Using this relationship, the output value,  $v(t)$ , is used to derive an estimate for the input,  $T_A$  (Racette and Lang, 2005).

It is of interest to develop a calibration approach that can detect gain abnormalities, and/or correct for slow drifts that affect the quality of the instrument measurements. To demonstrate the dynamic approach in terms of application appeal, Rivers and Boone (2014) used the dynamic approach to characterize a calibration target over time for a microwave radiometer. The data used for this example was collected during a calibration experiment that was conducted on the Millimeter-wave Imaging Radiometer (MIR) (Racette *et al.* 1995). The purpose of the experiment

was to validate predictions of radiometer calibration.

We extend the work of Rivers and Boone (2014) to the nonlinear model by examining a



**Figure 5:** Power output measurements:  $V_{cryo}$  (blue);  $V_{amb}$  (violet); and  $V_{warm}$  (red)

data set similar to one that one be created by a laboratory bench calibration experiment. The National Institute of Standards and Technology (N.I.S.T.) conducted a bench calibration experiment designed at studying calibration methods, due to intellectual property rights the data is not free to use, so with the aid of computers we simulated data with the attributes of the N.I.S.T. data. The data set consist of three temperature references and the corresponding three power output measurements collected over 1400 time periods. The references temperatures are as follows:

- $T_{cryo} = 84.3^\circ K$
- $T_{amb} = 296.2^\circ K$
- $T_{warm} = 300.7^\circ K$

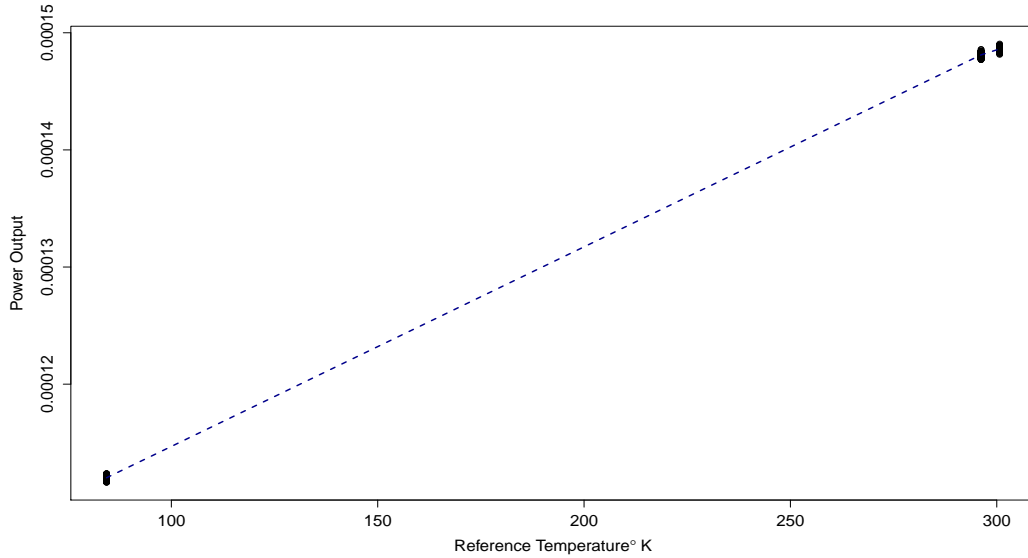
with summary statistics for the corresponding output measurements as

	$V_{cryo}$	$V_{amb}$	$V_{warm}$
$\bar{v}$	0.0001120096	0.0001481137	0.0001486190
$\sigma_v$	0.0000001280	0.0000001308	0.0000001236

The time series plots for the observed output measurements are provided in Figure 5.

In Figure 6 we plot the output measurements against the reference temperatures and show the calibration curve by the dotted line. One may suggest that a linear fit is appropriate but the experiment that inspired this data was created such that the quadratic term  $\beta_{2t}$  would be significant.

We demonstrate the dynamic calibration method on the simulated radiometer data by setting  $y_{0t} = 0.0001347169$  which corresponds to  $x_{0t} = 200^\circ K$  given a stable system without any drift in



**Figure 6:** 3-Point Calibration model used to detect nonlinearity. Nonlinear calibration curve (black dots connected by dotted red line) goes through three reference measurements collected over time.

the dynamic regression parameters. An assessment of the dynamic method is conducted at the end of the time series when  $t = 1400$ . After employing Algorithm 3 of the data we get a mean value for  $\hat{x}_{0t}$  across time as 199.5695 with a standard deviation  $\sigma_{x_{0t}} = 1.53135$ . It is of interest to now if the addition of references would improve the estimation process because the placement of the references in the experiment conducted at NIST may not completely capture the degree of the nonlinearity. In Figure 7 we added two additional reference measurements. We place the additional measurements as follow:

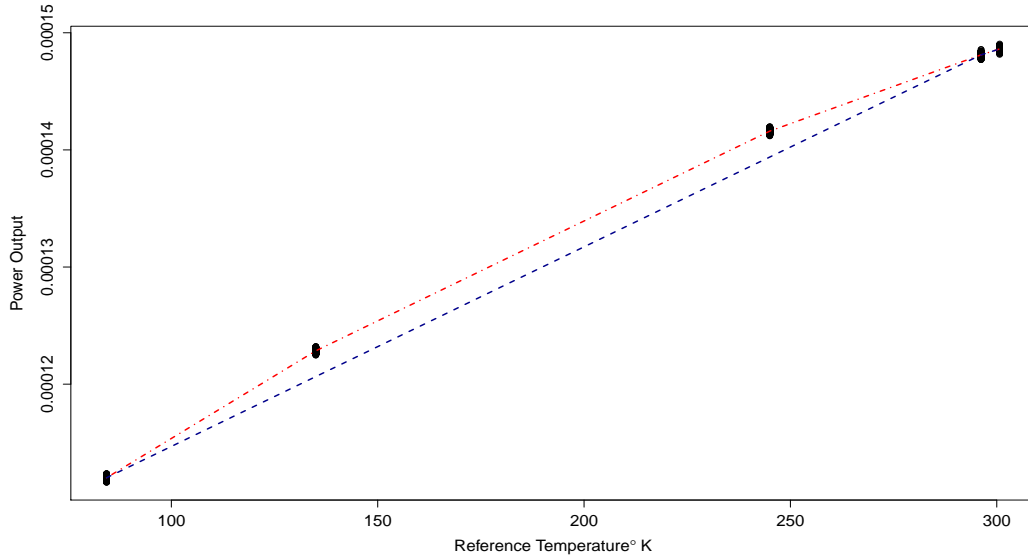
- $T_{135} = 135^\circ K$
- $T_{245} = 245^\circ K$

with the corresponding output measures listed in the table below

	$V_{135}$	$V_{245}$
$\bar{v}$	0.0001228344	0.0001415994
$\sigma_v$	0.0000001257	0.0000001233

The dynamic calibration method was conducted on the data set with the inclusion of the additional reference measurements yielding a mean value across time of  $\hat{x}_{0t} = 200.1122$  which is closer to the assumed true temperature measure of  $200^\circ K$  and a standard deviation of  $\sigma_{x_{0t}} = 0.4042387$ . In Table 8 we compare the performance criterions from Section 3 side-by-side for the data when using a 3-point calibration model and 5-point calibration model.

In Table 8 it is easy to see that the inclusion of the two reference measurements greatly improved the estimation of the calibration distributions across time. This is evident by the decrease in the mean square error. The interval width is shorter for the 5-point model because as stated before that standard deviation when the 2 measurements were added is significantly smaller than when not including them. The greatest improvement took place in the coverage probability



**Figure 7:** 5-Point Calibration model versus 3-Point Calibration model. The calibration curve for the 3-point model is shown by the dotted line and the calibration curve for the 5-point model is shown by the dash-dotted line.

**Table 8:** Summary of the 3-point calibration model versus the 5-point calibration model.

Model	MSE	IW	CP
3-Point	2.5130986	2.2480498	0.4450423
5-Point	0.1602064	1.8286866	0.9699140

(CP) because it nearly doubled by pacing the references between the endpoint measurements.

The results shown in this section indicate that, dynamic calibration can be used to determine concentrations, reference temperatures, or any other unknown measurement source where time period of calibration is significant. Until Rivers and Boone (2014) introduced the dynamic method statistical calibration has been considered from a static point of view. To add to the discussion of the dynamic Bayesian nonlinear calibration, next we consider a case where concern may arise when the observed measurement  $y_{0t}$  approaches the vertex. This sort of issue is considered to be problematic, thus we are concerned with how the dynamic method performs in this scenario.

## 5. FUTURE WORKS AND OTHER CONSIDERATIONS

Before concluding this work we would be remiss not to consider the case when the observed measurement  $y_{0t}$  approaches the vertex. We would like to understand how the calibration method will perform under such a condition. To understand the behavior of the vertex in a dynamic sense we look at the time-varying quadratic equation in vertex form,

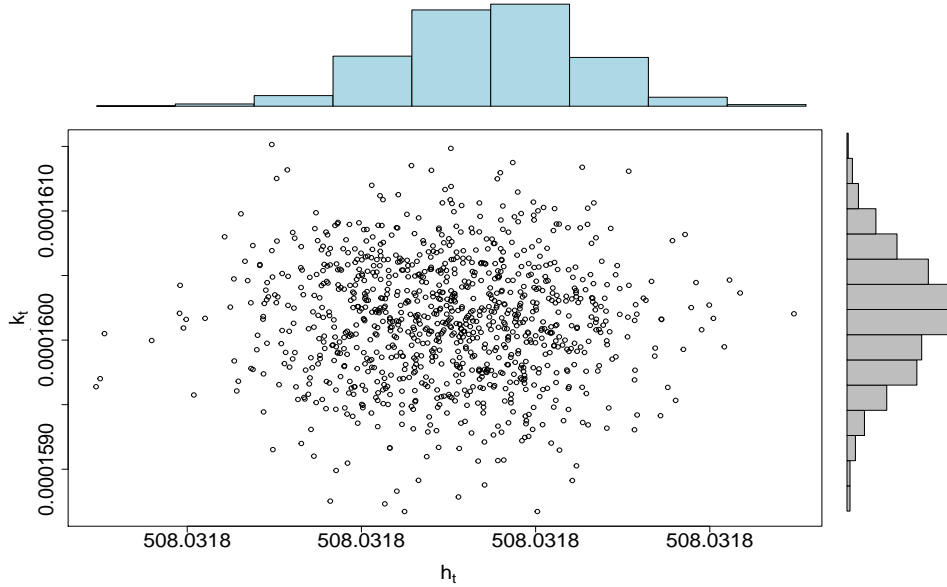
$$y_t = \hat{a}_t(x - \hat{h}_t)^2 + \hat{k}_t \quad (18)$$

where

$$\begin{aligned}\hat{a}_t &= \hat{\beta}_{2t}, \\ \hat{h}_t &= \frac{-\hat{\beta}_{1t}}{2\hat{\beta}_{2t}}, \\ \hat{k}_t &= \hat{\beta}_{0t} - \frac{\hat{\beta}_{1t}^2}{4\hat{\beta}_{2t}},\end{aligned}$$

and the time-varying vertex is  $(\hat{h}_t, \hat{k}_t)$ .

The estimated dynamic regression parameters are used to derive the joint distribution of



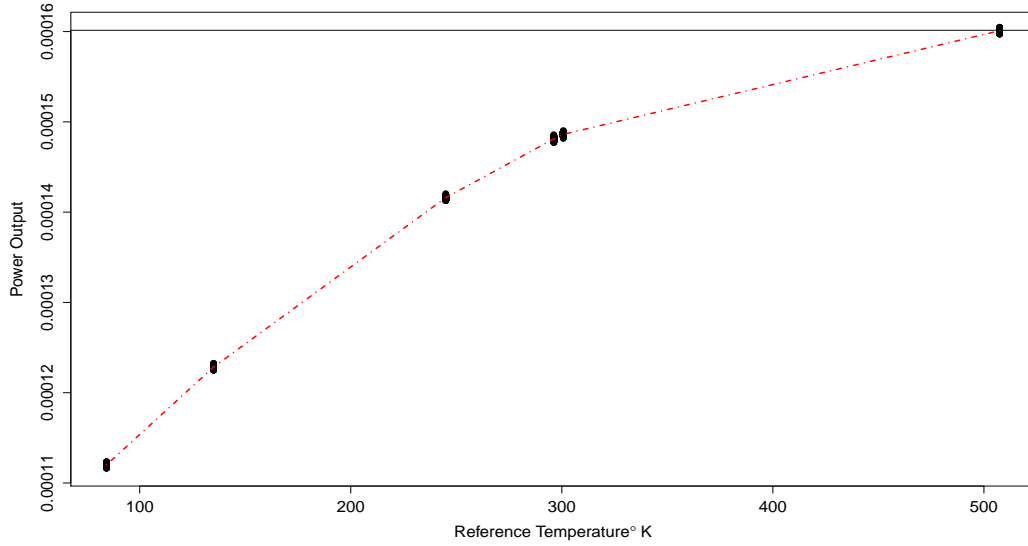
**Figure 8:** Joint distribution of  $(\hat{h}_t, \hat{k}_t)$  with marginal distributions of  $\hat{h}_t$  and  $\hat{k}_t$

$(\hat{h}_t, \hat{k}_t)$  given in Figure 8. So, we extend the simulated radiometer data previously used in Section 4 by placing a reference measurement at  $508^\circ\text{K}$ . A time series of observed power outputs  $y_{0t}$  is generated from the following quadratic equation:

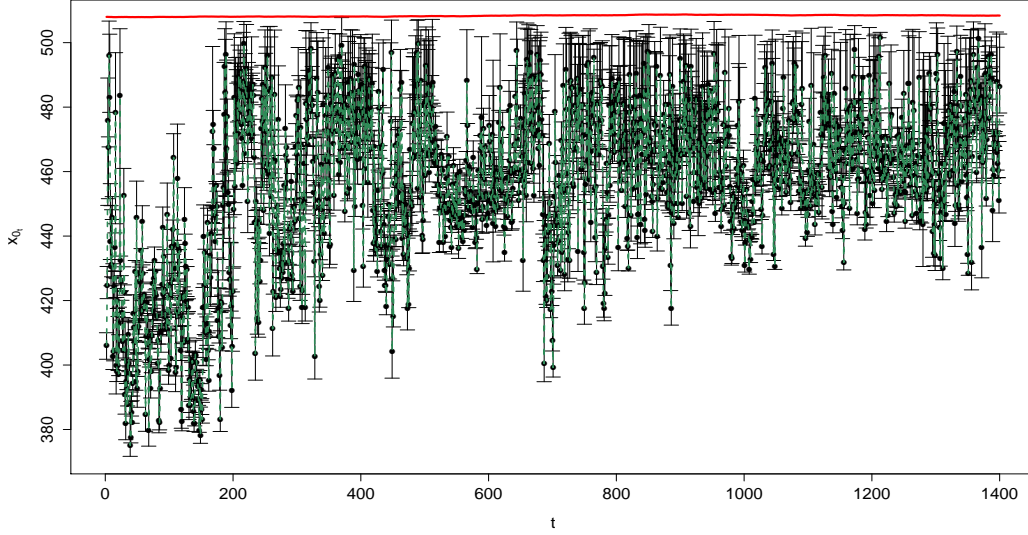
$$y_{0t} = \hat{\beta}_{0t} + \hat{\beta}_{1t}(508) + \hat{\beta}_{2t}(508)^2, \quad (19)$$

where  $\hat{\beta}_{0t}$ ,  $\hat{\beta}_{1t}$  and  $\hat{\beta}_{2t}$  are the estimates of the time dependent regression parameters. In Figure 9 the black horizontal line represents the possible maximum power output  $V_t$  given the vertex is  $(\hat{h}_t, \hat{k}_t)$  at time  $t$ .

We used dynamic calibration method described by Algorithm 3 to examine just how the method would perform at estimating the temperature value of  $508^\circ\text{K}$ . In Figure 10 we see that the method performs poorly from a mean squared error and coverage probability point of view. The method is restricted from deriving distributions that violate the quadratic behavior by not calculating credible intervals that go beyond any real possible value, meaning that the upper credible limit  $x_{0t}^U$  will be less than the true value for  $x_{0t}$  with  $P(x_{0t}^U < x_{0t}) = 1$ . Clearly, it is not advisable to perform a statistic calibration experiment so close to the vertex. Calibration, whether



**Figure 9:** Quadratic calibration model (dash-dotted line) with observation  $y_{0t}$  near vertex.

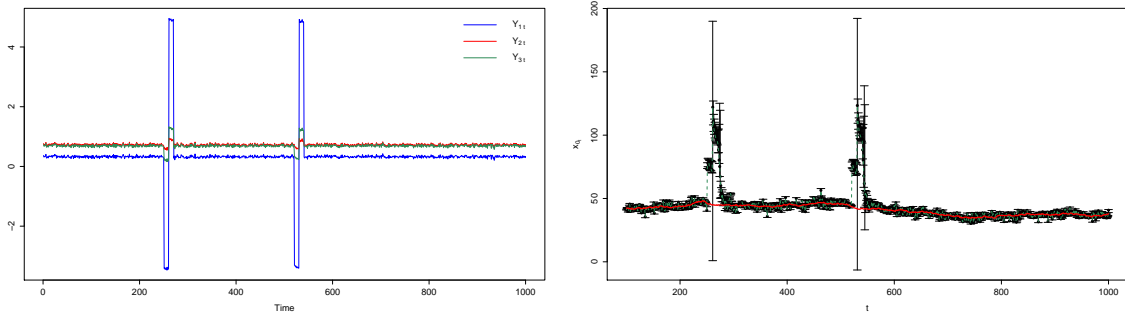


**Figure 10:** Quadratic calibration model (dash-dotted line) with observation  $y_{0t}$  near vertex.

static or dynamic is best conducted when the calibration target is near the center of the references domain space.

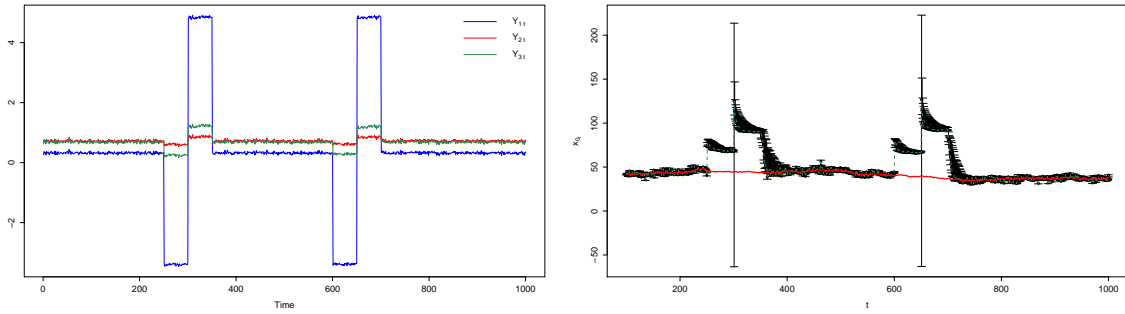
Finally, we consider an extreme case where some sort of shock is imposed with the system of the calibration device that causes drastic shifts in the observation data at random points in time. It is of interest to understand how the dynamic approach handles such a case. As before, we model





(a) Time series of 3 observation measurements with shocks at  $t = 250$  and  $t = 520$ . (b) Calibration distributions in the presence of system disturbances.

**Figure 11:** System disturbance for 20 time periods.



(a) Time series of 3 observation measurements with shocks at  $t = 250$  and  $t = 520$ . (b) Calibration distributions in the presence of system disturbances.

**Figure 12:** System disturbance for 100 time periods.

the dynamic system in vertex form,

$$y_t = \hat{s}_t(x - \hat{h}_t)^2 + \hat{k}_t \quad (20)$$

where

$$\begin{aligned} \hat{s}_t &= \hat{\beta}_{2t}\gamma_t, \\ \hat{h}_t &= \frac{-\hat{\beta}_{1t}}{2\hat{\beta}_{2t}}, \\ \hat{k}_t &= \hat{\beta}_{0t} - \frac{-\hat{\beta}_{1t}^2}{4\hat{\beta}_{2t}}, \end{aligned}$$

and the random multiplicative disturbance is denoted as  $\gamma_t$ . In Figure 11a we have three time series measurements, denoted as  $Y_{1t}$ ,  $Y_{2t}$ , and  $Y_{3t}$ , from a calibration experiment that is taken from the reference temperature measurements of  $20^\circ\text{K}$ ,  $90^\circ\text{K}$ , and  $100^\circ\text{K}$ . We see, that at times  $t = 250$  and  $t = 520$  a random shock has been imposed on the system thus causing mean-shifting glitches in the observation data. Each of the random shock periods last for exactly 20 time periods in Figure 11a. The disturbance to the system get incrementally greater as we move away from the

vertex, one would see larger fluctuations for the observations for the lower reference measurement than for the higher reference measurements. This supports the conclusions made by François et al. (2004). We see in Figure 11b that the calibration distributions credible intervals grow wider during the interrupted period as the dynamic method attempts to learn the behavior of the regression relationship. A consequent of the glitches is that the coverage probability will decrease and the interval widths will increase.

To get a better understand of the behavior of the method we extend the time frame of the system disturbance from 20 time periods to 100 time periods where half of the time the disturbance is negative and the other half it is positive. In Figures 12a and 12b we see that the disturbance in the system is translated in dynamic method. Eventhough the dynamic methods coverage probability is drastically lessened by the glitch it still outperforms the static method as the static quadratic calibration method is unable to derive estimates across time if there is instability in the parameters. Future work is to investigate the dynamic calibration methods from a multiple multivariate calibration point of view. In such settings we not only establish the inter-device calibration relationship to perform calibration over time of the particular instrument but we establish a dependency among the instruments as well.

## REFERENCES

- [1] Aitchison, J., and Dunsmore, I. R. (1975). *Statistical prediction analysis*. Cambridge University Press.
- [2] Albert, J. (2007). *Bayesian Computation with R* (Vol. 747389981). New York: Springer.
- [3] Berkson, J. (1969). Estimation of a Linear Function for a Calibration Line; Consideration of a Recent Proposal. *Technometrics*. **11**, 649-660.
- [4] Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference. *Journal of the Royal Statistical Society. Series B (Methodological)*. 113-147.
- [5] Bremer, J. C. (1979). Improvement of Scanning Radiometer Performance by Digital Reference Averaging. *Instrumentation and Measurement, IEEE Transactions on* 28.1: 46-54.
- [6] Brown, P. J. (1982). Multivariate calibration. *Journal of the Royal Statistical Society B*. **44**, 287-321.
- [7] Cervenka, P.O., and Massa, L. (1994). Research and Development Report: *Nonlinear Calibration of an Infrared Radiometer*. Paper presented at Workshop on Infrared Ship Signature Modeling (IRSSM), organized by the NATO Research Study Group on Maritime Infrared Target and Background Signatures, Measurements and Characteristics (RSG.5), Paris, France.
- [8] du Plessis, J. L., and van der Merwe, A. J. (1996). Bayesian calibration in the estimation of the age of rhinoceros. *Annals of the Institute of Statistical Mathematics*. **48**(1), 17-28.
- [9] Eisenhart, C. (1939). The interpretation of certain regression methods and their use in biological and industrial research. *Annals of Mathematical Statistics*. **10**, 162-186.
- [10] Eno, D. R. (1999). Noninformative Prior Bayesian Analysis for Statistical Calibration Problems. Doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.
- [11] Eno, D. R., and Ye, K. (2000). Bayesian reference prior analysis for polynomial calibration models. *Test*. **9**(1), 191-208.

- [12] Fong, Y., Wakefield, J., De Rosa, S., and Frahm, N. (2012). A Robust Bayesian Random Effects Model for Nonlinear Calibration Problems. *Biometrics*. **68**(4), 1103-1112.
- [13] François, N., Govaerts, B., and Boulanger, B. (2004). Optimal designs for inverse prediction in univariate nonlinear calibration models. *Chemometrics and intelligent laboratory systems*, **74**(2), 283-292.
- [14] Givens, G. H., and Hoeting, J. A. (2005). *Computational Statistics*. John Wiley and Sons.
- [15] Halperin, M. (1970). On Inverse Estimation in Linear Regression. *Technometrics*. **12**, 727-736.
- [16] Hibbert, D. B. (2006). The uncertainty of a result from a linear calibration. *Analyst*, **131**(12), 1273-1278.
- [17] Hoadley, B. (1970). A Bayesian look at Inverse Linear Regression. *Journal of the American Statistical Association*. **65**, 356-369.
- [18] Hunter, W., and Lamboy, W. (1981). A Bayesian Analysis of the Linear Calibration Problem. *Technometrics*. **23**, 323-328.
- [19] Imaoka, K., Kachi, M., Kasahara, M., Ito, N., Nakagawa, K., and Oki, T. (2010). Instrument Performance and Calibration of AMSR-E and AMSR2. *International Archives of the Photogrammetry, Remote Sensing and Special Information Science*. **38**(Part 8).
- [20] Kirkup, L., and Mulholland, M. (2004). Comparison of linear and nonlinear equations for univariate calibration. *Journal of Chromatography A*. **1029**(1), 1-11.
- [21] Krutchkoff, R. G. (1967). Classical and Inverse Regression Methods of Calibration. *Technometrics*. **9**, 425-439.
- [22] Lavagnini, I., and Magno, F. (2007). A statistical overview on univariate calibration, inverse regression, and detection limits: Application to gas chromatography/mass spectrometry technique. *Mass Spectrometry Reviews*. **26**(1), 1-18.
- [23] Lim, J. H., and Yun, H. I. (2010). Comparison of linear and nonlinear equation for the calibration of roxithromycin analysis using liquid chromatography/mass spectrometry. *Korean Journal of Veterinary Research*. **50**(1), 11-17.
- [24] Lundberg, E., and de Maré, J. (1980). Interval estimates in the spectroscopy calibration problem. *Scandinavian Journal of Statistics*. **7**(1) 40-42.
- [25] Lwin, T. and Maritz, J. S. (1982). An Analysis of the Linear Calibration Controversy from the Perspective of Compound Estimation. *Technometrics*. **24**, 235-242.
- [26] Martinelle, S. (1970). On the Choice of Regression in Linear Calibration. *Technometrics*. **12**(1), 157-161.
- [27] Merkle, W. (1983). Statistical methods in regression and calibration analysis of chromosome aberration data. *Radiation and environmental biophysics*. **21**(3), 217-233.
- [28] Naylor, J. C., and Smith, A. F. (1982). Applications of a method for the efficient computation of posterior distributions. *Applied Statistics*. 214-225.
- [29] Osborne, C. (1991). Statistical Calibration: A Review. *International Statistical Review*. **59**(3), 309-336.

- [30] Racine-Poon, A. (1988). A Bayesian approach to nonlinear calibration problems. *Journal of the American Statistical Association*. **83**(403), 650-656.
- [31] Rivers, D. L., and Boone, E. L. (2014). A Dynamic Approach to Linear Statistical Calibration with an Application in Microwave Radiometry. *arXiv preprint arXiv:1406.7863*.
- [32] Racette, P., and Lang, R. H. (2005). Radiometer Design Analysis Based Upon Measurement Uncertainty. *Radio Science*. **40**(5).
- [33] Rubin, D. B. (1987). The calculation of posterior distributions by data augmentation: Comment: A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: The SIR algorithm. *Journal of the American Statistical Association*, 543-546.
- [34] Smith, R. L., and Corbett, M. (1987). Measuring marathon courses: an application of statistical calibration theory. *Applied Statistics*. 283-295.
- [35] Smith, A. F., and Gelfand, A. E. (1992). Bayesian statistics without tears: a sampling-resampling perspective. *The American Statistician*, **46**(2), 84-88.
- [36] Tierney, L., and Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*. **81**(393), 82-86.
- [37] ter Braak, C. J. (1995). Nonlinear methods for multivariate statistical calibration and their use in palaeoecology: a comparison of inverse ( $k$ -nearest neighbours, partial least squares and weighted averaging partial least squares) and classical approaches. *Chemometrics and Intelligent Laboratory Systems*, **28**(1), 165-180.
- [38] Ulaby, F. T., Moore, R. K., and Fung, A. K. (1981). *Microwave Remote Sensing: Active and Passive, Vol. I: Fundamentals and Radiometry*. Addison-Wesley, Advanced Book Program, Reading, Massachusetts.
- [39] Weinreb, M. P., Hamilton, G., Brown, S., and Koczor, R. J. (1990). Nonlinearity corrections in calibration of Advanced Very High Resolution Radiometer infrared channels. *Journal of Geophysical Research: Oceans* (1978–2012). **95**(C5), 7381-7388.
- [40] West, M., Harrison, P. J., and Migon, H. S. (1985). Dynamic generalized linear models and Bayesian forecasting. *Journal of the American Statistical Association*, **80**(389), 73-83.
- [41] West, M. and Harrison, P.J. (1997). *Bayesian Forecasting and Dynamic Linear Models*, 2nd Edition. Springer-Verlag, New York.
- [42] Williams, E. J. (1969). A Note on Regression Methods in Calibration. *Technometrics*. **11**(1), 189-192.
- [43] Ziemer, R. R., and Strauss, D. (1978). A statistical approach to instrument calibration. USDA Forestry Service Technical Report.